

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)**ScienceDirect**

Procedia Computer Science 46 (2015) 499 – 506

**Procedia**  
Computer Science

International Conference on Information and Communication Technologies (ICICT 2014)

# Understanding the Usage of Online Forums as Learning Platforms

Smrithi Rekha V<sup>a,\*</sup>, Venkatapathy S<sup>b</sup><sup>a</sup>*Center for Research in Advanced Technologies for Education, Amrita Vishwa Vidyapeetham, Kerala, India*<sup>b</sup>*Gran Sasso Science Institute, L'Aquila, Italy*

---

## Abstract

Online forums provide an excellent platform for learning and connecting participants across the globe. However, only a few of them are suitable for learning. In this paper, we study the usage of StackOverflow (SO) forum through a survey and have presented our findings. The main objective of the study is to understand how satisfied users are with various aspects of the StackOverflow forum and how effective it is as a learning platform. The results reveal that StackOverflow may not support the learning needs of new learners. The specificity also discourages many users from frequently interacting on the forum.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)

**Keywords:** Online Forums; Online Learning Platforms; Stackoverflow Forum

---

## 1. Introduction

Online QA forums like Quora and StackOverflow (SO) are emerging as platforms that enable large users to interact on common topics of interest. SO is specifically for people discussing on computer programming languages like C, C++ and Python. Currently SO has more than three lakh users and ten lakh questions. SO follows a Question and Answer format and there are strict standards on the manner in which users can interact on the forum. Due to the volume of questions available on the forum, users are using SO not just as a QA forum but also as a learning

---

\* Corresponding author. Tel.: +91 9487006160 ; fax: +91 476 2896178  
E-mail address: [v\\_smrithirekha@cb.amrita.edu](mailto:v_smrithirekha@cb.amrita.edu)

platform for programming languages. Interested in understanding the usage of SO forum by users, we conducted a survey of SO forum users focusing on the frequency of usage, purpose of usage and their satisfaction with the forum as a learning platform. Answers to our research questions, we believe, will give us insights on the user's perception about the various aspects of the forum. These insights will help us offer recommendations to users and the forum on making the interactions more beneficial and of high quality. Our paper is organized as follows. In Section 2 we present background information on the forum and discuss various related work. In Section 3 we describe the survey methodology. The analysis of survey results is presented in Section 4. We present our main findings and recommendations in Section 5. We conclude in Section 6.

## 2. Background and Related Work

Discussion forums were conceptualized as early as 1980. The main objective was to create a platform for asynchronous communication among users. They are seen as online collaborative spaces where a diverse set of people converge and share knowledge. Due to this huge repository of knowledge contributed by people from all over the world, online forums can be effective sources of learning<sup>1, 2</sup>. They were also seen as enablers of blended and constructivist learning<sup>3</sup>.

StackOverflow is a Q&A forum for software professionals, students and programmers. Users can register for free and post their questions related to any programming concepts such as programming languages, databases, XML, etc. Each question can be accompanied with a set of tags that are used to categorize the question with other similar questions<sup>4</sup>. Tags are also used by a user to search for questions asked by the other users. Users can get reputation scores based on their contribution and interaction on the forum. Users and questions also have badges attached to them to indicate their quality.

### 2.1. Related Work

Over the years, a plethora of research has been done on the effectiveness of online forums as learning platforms. Thomas in<sup>5</sup> has studied the learning outcomes and patterns of interactions on online discussion forums. With respect to the StackOverflow forum, Treude et al tried in<sup>6</sup> to analyze the role of Q&A websites in the software development. The authors of<sup>7, 8</sup> have analyzed SO from the perspective of the type of questions on the forum. In<sup>9</sup>, Wang et al published a study on developers' interaction in SO. Nasehi et al in<sup>10</sup> tried to address the question of what makes a good code example by categorizing the questions of SO. One of their main findings is that a question with a concise code accompanied with a good explanation can make it to a good code example. In<sup>11</sup> the authors studied the correlation between the usage of classes of Android, and questions asked related to that respective classes in SO.

Though there have been many studies in the past to study the different perspectives of SO as a community based Question Answering site, as far as we know, no research has been done to identify the effectiveness of SO as a learning platform. The results presented in<sup>12</sup> come close to our study where they tried to analyze the impact of Q&A of knowledge sharing in open source communities. In our previous work<sup>13</sup>, we have done a classification of users using X-Means clustering algorithm. In<sup>16</sup>, we have implemented an auto-tagging system for the questions of SO forum.

## 3. Methodology

The main purpose of the study was to understand the effectiveness of SO as a learning platform. To realise this objective, we studied how people use the forum generally, their interactions on the forum and how much they learnt through the forum. Such an understanding, combined with insights from our previous work on machine learning based user classification, will help us offer recommendations on making the interactions more fruitful for the users. The main research questions that guided our study are:

- RQ1: What is the frequency and purpose of using SO forum?
- RQ2: What is the satisfaction of users with various aspects of the forum?
- RQ3: How effective is SO as a learning platform?

### 3.1. Survey Design

A questionnaire was created based on these research questions. Since SO is specifically about programming, we chose the sample population from people who were involved with some form of programming on a regular basis. The questionnaire had 29 questions including branching questions. Except for the personal profile based questions, all the questions were objective-type questions involving no descriptive types. For easy analysis the questions have been numbered sequentially. Q1 to Q6 are related to the *personal and professional profile* of the respondents. Q7 to Q14 are about the *generic usage of the SO forum*. Q15 to Q19 are for those who have *posted questions on the forum* and Q20 to Q25 are for those who *have answered questions on the forum*. Q26 is a rating question, where respondents rate their *satisfaction levels* with various aspects of the forum on a scale of 1-5. Q27 and Q28 are related to the user's *perception of SO as a learning platform*. Q29 is related to *user classification*.

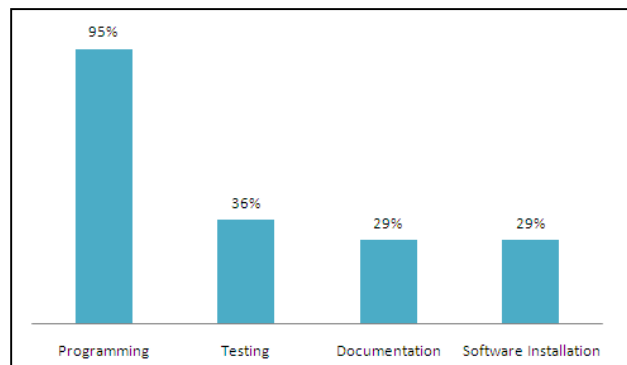


Fig. 1. Job Responsibilities of Respondents

### 3.2. Data Analysis Procedure

The responses to the questions were analyzed in the following manner

- The responses were collected in a spreadsheet
- The questions were numbered and then a question-wise analysis was performed
- Frequency of responses were calculated for objective type questions
- T-Tests were performed on Likert Scale questions
- Summary of main findings was recorded

## 4. Analysis of Results

### 4.1. Participant Profile

A total of 57 people had participated in the survey. Nearly 57% of the respondents are from the Software Industry involved with roles as Software Engineers, Developers, QA Engineers, Data Analyst and Programmers. The rest of the respondents are students, researchers, faculty and consultants. The spread of respondent profiles is shown in Figure 1. 95% of the respondents are involved with Programming, 36% do testing, 29% of them have indicated that documentation and Software Installation cum support is part of their jobs/daily work.

#### 4.2. General Usage of the Forum

In this section, the participants were asked about their general usage of the forum. On an average, respondents

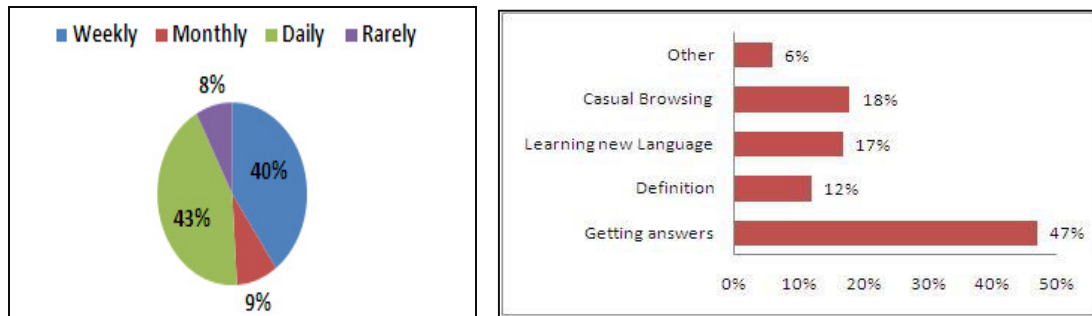


Fig. 2. Stackoverflow Usage frequency and Purpose

have known and used SO forum for about 34 months. About 40% of them use it weekly, 43% of them use it daily, about 9% of them use it monthly and only 7% of the respondents use it rarely. The various purposes for which they use the forum include (percentage of respondents using it for a particular purpose is also indicated here):

- Getting solutions to programming problems they encounter: 51%
- To know definitions of programming concepts: 13%
- To learn new programming concepts: 18%
- Casual Browsing of Questions and Answers: 19%
- Other purposes like getting suggestions on learning material or learning a new programming language : 7%

It is very interesting to note that more than 50% of the users use search engines like google while encountering programming issues and then proceed to SO to find solutions. About 70% of the respondents are registered users of the forum. This means that they enjoy several privileges on the forum like posting questions, answering them, editing tags etc. The remaining respondents, who are non-registered users, can browse through the forum but cannot actively participate in the forum. More than 50% of the respondents have either posted questions or answered questions or voted on the forum. Though there are a huge number of users with very high reputation score, very few of the respondents actively follow the questions and answers posted by these reputed users unlike other social networks where users are able to like or follow reputed users' pages.

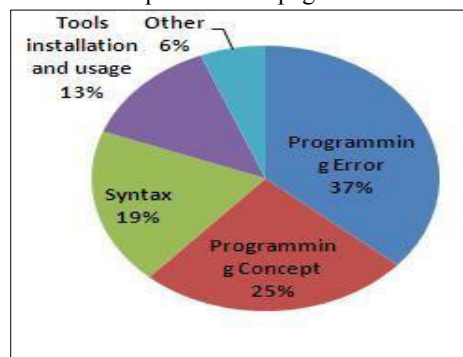


Fig. 3. Type of Questions Posted

### 4.3. Posting Questions

As mentioned earlier, more than 50% of the respondents have posted questions on the forum. The different types of questions that they have posted include a) Programming error related b) Programming concepts related c) Programming language syntax related d) Software tools installation and usage e) Other purposes like how to questions. The distribution of question types posted by respondents is shown in Figure 3. Though most respondents have indicated that they post questions, the frequency of posting seems to be less. About 90% of users post questions less frequently. 50% of them have received badges for their questions. The different types of badges that users have received are *Popular Question*, *Scholar*, *Tumbleweed*, *Favorite Question*, *Famous Question*.

### 4.4. Posting Answers

It is surprising to note that more than half of the respondents *rarely* answer questions on the forum and those that answer, respond to questions related to specific programming languages. About 88% of them have received votes and badges on their questions. Some of the badges that respondents have received include *Excavator*, *Commentator*, *Supporter*, *Self-Learner*.

### 4.5. Likert Scale Questions

#### 4.5.1. Satisfaction with SO

We had eight likert scale questions where participants had to rate their satisfaction levels with various aspects of the forum. We did a question-wise statistical analysis the results of which are shown in Figure 4. We framed the null hypothesis for each question from Q26 to Q33. For most of the questions, the average rating is above 3.0. The results of the statistical analysis are as follows:

- Q26.H0:Solutions in SO forum are not satisfactory : Rejected
- Q27.H0:Accuracy of solutions in SO are not satisfactory: Rejected
- Q28.H0:Networking Possibilities in SO are not satisfactory: Accepted
- Q29.H0:Level of detail in answers are not satisfactory: Rejected
- Q30.H0:Breadth and depth of topics covered are not satisfactory : Rejected
- Q31.H0:Sufficiency question tags in the form are not satisfactory : Rejected
- Q32.H0:Accuracy of questions tags are not satisfactory: Rejected
- Q33.H0:Response time is not satisfactory: Rejected

We see that users are satisfied with various aspects of SO. However, the t-test for Q28 shows that users are not satisfied with the networking possibilities on the forum. Though it's a forum which has millions of users, it is surprising to note that people use the forum more to get solutions to programming problems than interact with each other or form network of connections. While it may be argued that it is the very purpose of the forum, we cannot ignore that in such a large forum people network very less or track specific users. It would be useful to study the reasons behind this and compare with other forums like Quora which is also a QA forum but people are encouraged to interact.

<pre>data: Dataset\$Q26 t = 11.7838, df = 54, p-value &lt; 2.2e-16 alternative hypothesis: true mean is not equal to 3 95 percent confidence interval:  3.995833 4.404167 sample estimates: mean of x  4.2</pre>	<pre>data: Dataset\$Q27 t = 13.6751, df = 54, p-value &lt; 2.2e-16 alternative hypothesis: true mean is not equal to 3 95 percent confidence interval:  4.008554 4.355082 sample estimates: mean of x  4.181818</pre>	<pre>data: Dataset\$Q28 t = -0.531, df = 54, p- value = 0.5976 alternative hypothesis: true mean is not equal to 3 95 percent confidence interval:  2.652690 3.201855 sample estimates: mean of x  2.927273</pre>
--	---	---

Fig. 4. T-Test Snapshots for Likert-Scale Questions

#### 4.5.2. Satisfaction with SO

Participants were asked to rate the usefulness of various aspects of SO forum. For each question from Q34 to Q39, we formed the null hypothesis and tested them using one-sample t-test. The results of the analysis are:

- Q34.H0:SO is not very useful for self-learning of programming languages: Rejected
- Q35.H0:SO is not very useful for clarifying doubts related to programming languages: Rejected
- Q36.H0:SO is not very useful in gaining knowledge about computer science: Rejected
- Q37.H0:SO is not very useful in enhancing coding capabilities: Rejected
- Q38.H0:SO is not very useful in Enhancing Forum participation and Communicative abilities : Accepted
- Q39.H0:SO is not very useful for conducting research on online forums: Rejected

Very similar to the previous sets of questions we note that participants find SO to be less useful when it comes to enhancing their forum participation and communicative abilities. Respondents have indicated that once they find solution to their issues, they stop there and do not conduct a further search for solutions. A snapshot of the t-test results for likert-scale based questions is shown in Figure 4.

#### 4.6. Self-Categorization

In our previous work, we had algorithmically classified the users. We were interested in knowing how users classify themselves depending on their participation in the forum. We had defined five different categories.

- **Solution Finder:** One who participates in the forum to just find solutions
- **Naive Users:** Users who post very less number or no queries or answers ,has a low lifetime in SO and owns very few or no badges
- **Surpassing Users:** Users who post very less number of queries or answers, stay with SO for more number of days than naive users and own a few badges
- **Experts:** Users who post good number of queries and answers with many Up Votes , stay with SO for a long time and own a variety of large badges
- **Outshiners:** Users who post maximum number of queries and answers with most Up Votes , stay with SO for a really long time and own a variety of maximum Badges

Contradictory to our automated classification of users, most respondents (60%) have classified themselves as *Solution Finders* and few of them as *Surpassing Users*. This very well correlated with the fact that not all respondents who post and answer questions and continue to actively participate in the forum and just remain as dormant users of the forum.

## 5. Analysis of Results

### 5.1. Main Findings

Based on the analysis of survey data, our main findings are as follows:

- Participants use the forum more for getting answers to specific problems in Computer Programming and less for generic discussion on programming topics
- Survey participants post and answer questions less frequently but use it more for viewing solutions when they encounter problems. This is true for both participants from India and abroad.
- Users are satisfied with various aspects of the forum like accuracy of solutions, sufficiency of tags etc but when it comes to networking possibilities, users are less satisfied.
- Unlike Quora, which encourages new learners, SO is for well-informed users who already have good exposure to Computer Programming.
- Participants do not find SO useful for enhancing their communication and forum participation skills
- Very rarely do users directly use the SO forum i.e they web search for solutions and land up in SO as their first choice
- Most users categorize themselves as Solution Finders and not active participants on the forum.
- The solution matters more than the person who has provided the solution.

### 5.2. Recommendations

Though SO carries a high reputation, we see from this study that most respondents use the forum but do not actively participate in the forum. We also clearly see that SO does not satisfy the communication and discussion needs of users. Due to high level of specificity, SO may have less value as a full-fledged learning platform where users can expand their knowledge of various topics. Hence SO may be a forum that enhances the cognitive skills of users but not cater to the interaction and social requirements of users participating in online forums. Some recommendations that we offer to increase the active participation of users and enhance the learning potential of the forum are:

- Motivate new comers to the forum by encouraging broad questions and then guiding them on asking specific questions. This will help to enhance the utility of SO as a learning platform
- Have separate pages where a pool of resources like websites, book names etc are maintained for various programming languages
- Enhancing the interaction levels of users on the forum by bridging the gap between users who post questions and those that answer them. This could be done by allowing users to follow each other similar to Quora and Facebook. This will help enhance the interaction dimension of the forum as mentioned in <sup>14, 15</sup>
- Enhancing user participation by featuring top users. Currently the forum is more focused on the questions. It may be useful to focus on users by promoting high performing users on the home page of the forum or on the programming language page. This will also bring help connect users with each other.
- Encouraging users to provide more detailed answers by pointing to additional resources may help to enhance the knowledge base available on SO. This will reduce the granularity of certain questions on the forum, making it more suitable for learners and strengthening the pedagogical aspect of the forum

## 6. Conclusion

In this paper we have presented a detailed analysis of how people use the SO forum and how effective it is as a learning platform. While SO is suited for users having a reasonably good knowledge of Computer Programming, certain aspects of the forum may be adapted to suit new comers. This will enhance the quality of discussions on the forum and also push more users from just being naïve users on the forum to active participants in the forum. This will also enrich the utility of the forum as a learning platform. As future work we propose to do topic modeling on



the questions on the forum to understand the quality of questions. We also propose to do a detailed comparative analysis of generic QA forums like Quora and specific forums like Stackoverflow.

## References

1. G. Saade and Q. Huang, Meaningful Learning In Discussion Forums: Towards Discourse Analysis, *Issues in Informing Science and Information Technology*; vol. 6, p. 87–99, 2009.
2. J. Luca and C. McLoughlin, Using Online Forums To Support A Community Of Learning, in *World Conference on Educational Multimedia, Hypermedia and Telecommunications*; vol. 2004, no. 1, 2004, p. 1468– 1474.
3. P. Valiathan, Blended Learning Models, *Learning circuits*, 2002.
4. A. Saha, R. Saha, and K. Schneider, A Discriminative Model Approach For Suggesting Tags Automatically For Stack Overflow Questions, in *Mining Software Repositories (MSR), 2013 10th IEEE Working Conference on*; May 2013, p. 73–76.
5. M. J. Thomas, Learning Within Incoherent Structures: The Space of Online Discussion Forums, *Journal of Computer Assisted Learning*; vol. 18, no. 3, p. 351–366, 2002.
6. C. Treude, O. Barzilay, and M.-A. Storey, How Do Programmers Ask and Answer Questions on the Web, in *Software Engineering (ICSE), 2011 33rd International Conference on*; IEEE, 2011, p. 804–807.
7. Allamanis, M., & Sutton, C, Why, when, and what: analyzing stack overflow questions by topic, type, and code, In *Proceedings of the 10th Working Conference on Mining Software Repositories*; (p. 53-56), IEEE Press.
8. D. Correa and A. Sureka, Fit or Unfit: Analysis and Prediction of ‘Closed Questions’ on Stack Overflow, in *Proceedings of the first ACM conference on Online social networks*; ACM, 2013, p. 201–212.
9. S. Wang, D. Lo, and L. Jiang, An Empirical Study on Developer Interactions in Stackoverflow, in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*; ACM, 2013, pp. 1019–1024.
10. S. M. Nasehi, J. Sillito, F. Maurer, and C. Burns, What Makes A Good Code Example?: A Study of Programming Q&A In Stackoverflow, in *Software Maintenance (ICSM), 2012 28th IEEE International Conference on*; IEEE 2012, p. 25–34.
11. D. Kavalier, D. Posnett, C. Gibler, H. Chen, P. Devanbu, and V. Filkov, Using And Asking: Apis Used In The Android Market And Asked About In Stackoverflow, in *Social Informatics*; Springer, 2013, p. 405–418.
12. Sowe, Sulayman K., Ioannis Stamelos, and Lefteris Angelis. Understanding knowledge sharing activities in free/open source software projects: An empirical study. *Journal of Systems and Software*; 81.3 (2008): 431-446.
13. P. B. S. Smrithi Rekha, Anusha J, A Machine Learning Approach to Cluster the Users of Stack Overflow Forum, in *Intl Conference on Artificial Intelligence and Evolutionary Algorithms in Engineering Systems*; 2014.
14. F. Henri, Computer Conferencing and Content Analysis, in *Collaborative learning through computer conferencing*; Springer, 1992, p. 117–136.
15. H. Kanuka and T. Anderson, Online Social Interchange, Discord, and Knowledge Construction, *International Journal of E-Learning & Distance Education*; vol. 13, no. 1, pp. 57–74, 2007.
16. V. Smrithi Rekha, N. Divya, and P. Sivakumar Bagavathi. 2014. A Hybrid Auto-Tagging System for Stackoverflow Forum Questions. In *Proceedings of the 2014 International Conference on Interdisciplinary Advances in Applied Computing (ICONIAAC '14)*; ACM, New York, NY, USA